

An extension of Furstenberg's structure theorem for Noetherian modules and multiple recurrence theorems II

Xiongping Dai

Department of Mathematics, Nanjing University, Nanjing 210093, People's Republic of China

Abstract

Using a recent Furstenberg structure theorem, we obtain a quantitative multiple recurrence theorem relative to any locally compact second countable Noetherian module over a syndetic ring.

Keywords: Furstenberg theory · Noetherian module.

2010 MSC: Primary 37A15, 37A45 Secondary 37B20, 37P99, 22F05

Introduction

This paper will be devoted to studying the Furstenberg multiple recurrence of dynamical systems induced by any locally compact second countable Noetherian module over a syndetic ring acting on a standard Borel probability space, as a subsequent work of [6].

First of all, we recall that by an “lcscN” R -module G over a “syndetic” ring $(R, +, \cdot)$, we mean that G and $(R, +)$ both are locally compact second countable Hausdorff commutative groups such that:

- The operation $(t, g) \mapsto t \cdot g$ is continuous from $R \times G$ to G .
- G is a Noetherian R -module: for every sequence $G_0 \subseteq G_1 \subseteq G_2 \subseteq \cdots$ of R -submodules of G , we have $G_n = G_{n+1}$ as n sufficiently large.
- R is syndetic: $\forall t \neq 0$, Rt is syndetic in the sense that one can find a compact subset K of R with $K + Rt = R$.

See [6]. Clearly, $(\mathbb{Z}^n, +)$ over \mathbb{Z} , $(\mathbb{Q}^n, +)$ over \mathbb{Q} , and $(\mathbb{R}^n, +)$ over \mathbb{R} all are lcscN modules over syndetic rings. Moreover, the p -adic integer (syndetic) ring \mathbb{Z}_p and the (syndetic) p -adic number field \mathbb{Q}_p both are lcscN as modules over themselves. See [14] for more examples.

Let G be an lcsc R -module and let (X, \mathcal{X}, μ) be a standard Borel probability space. We will consider a measure-preserving Borel G -action dynamical system:

$$T: G \times X \rightarrow X \quad \text{or write} \quad G \curvearrowright_T X$$

Email address: xpdai@nju.edu.cn (Xiongping Dai)

where $T_g: X \rightarrow X$ is μ -preserving for each $g \in G$ and the G -action map $T: (g, x) \mapsto T_g(x)$ is jointly measurable from $G \times X$ to X . For brevity, we will write

$$T(t \cdot g, x) = T_{t \cdot g}(x) = T_g^t(x) = g^t x, \quad \forall t \in \mathbb{R} \text{ and } g \in G,$$

if no confusion. Then given any $g \neq e$ the identity element of G ,

$$T_g: \mathbb{R} \times X \rightarrow X \quad \text{by } (t, x) \mapsto T_g^t(x) = g^t x$$

defines a new μ -preserving Borel \mathbb{R} -system.

A metatheorem of dynamical theory states that whenever the underlying space of the dynamical system is appropriately bounded in the sense of topology or measure theory, the orbits of the motion will necessarily exhibit some form of recurrence, or return close to their initial position. The first precise theorem of this kind was formulated by H. Poincaré 1890 [17].

A quantitative recurrence theorem is due to A. Y. Khintchine as follows. Let \mathbb{R} be equipped with the standard Euclidean topology. Then Khintchine's Recurrence Theorem says that:

- If $T: \mathbb{R} \times X \rightarrow X$ is a C^0 -flow on a compact metric space X preserving a Borel probability measure μ , then for any $E \in \mathcal{B}_X$ with $\mu(E) > 0$ and $\varepsilon > 0$, the set

$$\{t \in \mathbb{R} \mid \mu(E \cap T^{-t}E) > \mu(E)^2 - \varepsilon\}$$

is relatively dense in \mathbb{R} (cf. [13, 15]).

Poincaré's Recurrence Theorem says that $\mu(E \cap T^{-t}E) > 0$ for infinitely many values of $t \in \mathbb{R}$; but the Khintchine Recurrence Theorem says that $\mu(E \cap T^{-t}E)$ is 'large' for 'many' values of $t \in \mathbb{R}$. Khintchine's theorem has been extended from \mathbb{R} -actions to σ -compact amenable group actions in [5]. Notably V. Bergelson, B. Host and B. Kra have recently derived in 2005 [2] the following multiple version of Khintchine's theorem for any μ -ergodic \mathbb{Z} -action $T: \mathbb{Z} \times X \rightarrow X$:

- Each of the following two returning-time sets, for any $E \in \mathcal{B}_X$ with $\mu(E) > 0$ and $\varepsilon > 0$, is relatively dense in \mathbb{Z} :

$$\{n \in \mathbb{Z} \mid \mu(E \cap T^{-n}E \cap T^{-2n}E) > \mu(E)^3 - \varepsilon\}$$

and

$$\{n \in \mathbb{Z} \mid \mu(E \cap T^{-n}E \cap T^{-2n}E \cap T^{-3n}E) > \mu(E)^4 - \varepsilon\}.$$

See [2, Theorem 1.2].

- Also see [8, 18] for polynomials version of the above multiple Khintchine theorem of Bergelson et al.

It should be noted that Bergelson et al have proven that the multiple version of Khintchine's theorem does not need to hold for *non-ergodic* \mathbb{Z} -actions (even for the 2-multiple case $\{n, 2n\}$, cf. [2, Theorems 2.1]) and does not need to be true for more *higher multiple* case $\{n, 2n, 3n, 4n\}$, cf. [2, Theorem 1.3]) by constructing counterexamples based on F.A. Behrend's and I. Ruzsa's combinatorial results.

However, by using completely different technique framework, we can obtain the following theorem, which claims any measure-preserving standard Borel G -system always has a nontrivial factor that possesses the multiple Khintchine recurrence. That is the following

Theorem 0.1 (Multiple Khintchine Recurrence). *Let G be an lcscN R -module over a syndetic ring $(R, +, \cdot)$ and let (X, \mathcal{X}, μ) be a nontrivial standard Borel G -space. Then there exists a G -invariant σ -subalgebra \mathcal{X}_1 of \mathcal{X} with $\{\emptyset, X\} \neq \mathcal{X}_1 \pmod{0}$ such that for any $A \in \mathcal{X}_1$ with $\mu(A) > 0$, any $g_1, \dots, g_l \in G$, $l \geq 2$, and any $\varepsilon > 0$, the set of returning times of big-hitting*

$$N_{g_1, \dots, g_l}(A, \varepsilon) = \left\{ t \in R \mid \int_X g_1^t 1_A \cdots g_l^t 1_A d\mu > \mu(A)^l - \varepsilon \right\},$$

is of positive lower Banach density in $(R, +)$; i.e.,

$$\liminf_{n \rightarrow \infty} \frac{|F_n \cap N_{g_1, \dots, g_l}(A, \varepsilon)|}{|F_n|} > 0$$

over any weak Følner sequence $\{F_n\}_1^\infty$ in $(R, +)$. Consequently it is syndetic in $(R, +)$.

Here a sequence of nonnull compact subsets F_n of positive Haar measure in $(R, +)$ with a Haar measure $|\cdot|$, which is such that

$$\lim_{n \rightarrow \infty} \frac{|(r + F_n) \triangle F_n|}{|F_n|} = 0 \quad \forall r \in R,$$

is called a *weak Følner sequence* in $(R, +)$.¹

Outline of the proof of Theorem 0.1 For example, $(\mathbb{R}^n, +)$ is an lcscN \mathbb{R} -module but it is not a free abelian group. So our Theorem 0.1 is already beyond the framework of Furstenberg-Katznelson [11, 10]. To prove Theorem 0.1 in the (not necessarily ergodic) probabilistic settings, our main tool is the following recent structure theorem.

Theorem 0.2 (Furstenberg Structure Theorem [6]). *Let G be an lcscN R -module over a syndetic ring $(R, +, \cdot)$. Then for any nontrivial standard Borel G -system $X = (X, \mathcal{X}, \mu, G)$, there exists an ordinal η and a system of G -factors $\{\pi_\xi: X \rightarrow X_\xi\}_{\xi \leq \eta}$ such that:*

- (a) X_0 is the one-point G -system and $X_\eta = X \pmod{0}$.
- (b) If $0 \leq \theta < \xi \leq \eta$, then there is a factor G -map $\pi_{\xi, \theta}: X_\xi \rightarrow X_\theta$ with $\pi_\theta = \pi_{\xi, \theta} \circ \pi_\xi$.
- (c) For each ordinal ξ with $0 \leq \xi < \eta$, $\pi_{\xi+1, \xi}: X_{\xi+1} \rightarrow X_\xi$ is a nontrivial “primitive” extension.
- (d) If ξ is a limit ordinal $\leq \eta$, then $X_\xi = \varprojlim_{\theta < \xi} X_\theta$.

Moreover, the intermediate factors are of the form

$$X_\xi = (X, \mathcal{X}_\xi, \mu, G), \quad \pi_\xi = Id_X \quad \text{and} \quad \pi_{\xi+1, \xi} = Id_X \quad (0 < \xi < \eta).$$

Here we refer to

$$X \rightarrow X_\eta \rightarrow \cdots \rightarrow X_{\xi+1} \rightarrow X_\xi \rightarrow \cdots \rightarrow X_1 \rightarrow X_0$$

a “Furstenberg factors chain” of X .

¹Here and in the future, unlike the case in literature, we neither need to require

$$\lim_{n \rightarrow \infty} \frac{|(K + F_n) \triangle F_n|}{|F_n|} = 0 \quad \forall K \subset R \text{ compact}$$

nor other regularities such as Tempelman condition or Shulman condition (cf. [7]).

Since X_0 is just the one-point G -system, then the first intermediate factor $X_1 = (X, \mathcal{X}_1, \mu, G)$ in the Furstenberg factors chain of X is such that $(X, \mathcal{X}_1, \mu, G_c)$ is compact and $(X, \mathcal{X}_1, \mu, G_w)$ is totally weak-mixing, where $G = G_c \times G_w$ is some direct product of two R -submodules of G . Then we only need to show X_1 is a Kh -system. See Proposition 4.1 in §4.

The remainder of this paper will be organized as follows:

Contents

1	A van der Corput-type lemma	4
2	Inverse limits of Kh-systems	7
3	Kh-property of totally relatively weak-mixing extensions	8
3.1	Relatively weak-mixing extensions	9
3.2	Kh -property of totally relatively weak-mixing extensions	11
4	Kh-property of chaotic G-systems	14
4.1	Compact G -system	14
4.2	Chaotic G -system	14
4.3	Proof of Theorem 0.1	16

Acknowledgments

Finally, the author is deeply grateful to Professor Hillel Furstenberg for many helpful suggestions, comments, and carefully checking the details of the original manuscript.

1. A van der Corput-type lemma

In this section we will prove a technical tool—a van der Corput-type lemma—for proving the Kh -property of a totally weak-mixing system in §3.

- Let $(R, +)$ be an lcsc abelian group with the zero element \mathbf{o} and with a fixed Haar measure $|\cdot|$ or dt such that $|R| = \infty$ in the sequel of this section. Here R is not necessarily a topological ring.

We can then take a weak Følner sequence $\{F_n\}_1^\infty$ in $(R, +)$; that is, $\{F_n\}_1^\infty$ is a sequence of compact subsets of positive Haar measure of R satisfying the so-called *weak Følner condition* [16]:

$$\lim_{n \rightarrow \infty} \frac{|(r + F_n) \triangle F_n|}{|F_n|} = 0 \quad \forall r \in R.$$

Given any measurable subset S of R , its *upper density over* $\{F_n\}_1^\infty$ is defined as follows:

$$D^*(S) = \limsup_{n \rightarrow \infty} \frac{|S \cap F_n|}{|F_n|}.$$

If $D^*(S) = 0$ over $\{F_n\}_1^\infty$ then we say that S has *density 0 over* $\{F_n\}_1^\infty$. By the weak Følner condition we note that over the same weak Følner sequence $\{F_n\}_1^\infty$ in $(R, +)$,

$$D^*(S) = D^*(S + r) \quad \forall r \in R.$$

We can similarly define the *lower density* $D_*(S)$ of S over $\{F_n\}_1^\infty$ by replacing the \limsup by \liminf and further define *density* $D(S)$ over $\{F_n\}_1^\infty$.

We shall speak of convergence in density over $\{F_n\}_1^\infty$ for a family of points but for a subfamily of 0-density. The following is the precise definition.

Definition 1.1. Let $(R, +)$ be an lcsc abelian group and $\{F_n\}_1^\infty$ a weak Følner sequence in it.

- (1) For a Borel measurable function $x : R \rightarrow X$ from R into a topological space X , we say that x *converges in density over $\{F_n\}_1^\infty$ to a point $x \in X$* if for every neighborhood V of x in X , $x_t \in V$ but for a set of t of $\{F_n\}_1^\infty$ -density 0. We write

$$\text{D-lim}_{t \in R} x_t = x \quad \text{over } \{F_n\}_1^\infty.$$

- (2) Let $a : R \rightarrow \mathbb{R}$ be a measurable function. The *ess.l.u.b.* (possibly infinite)

$$\text{D-lim sup}_{t \in R} a_t := \inf \{c : D^*(\{t \in R \mid a_t > c\}) = 0 \text{ over } \{F_n\}_1^\infty\}$$

or equivalently

$$\text{D-lim sup}_{t \in R} a_t = \inf \{c : D(\{t \in R \mid a_t \leq c\}) = 1 \text{ over } \{F_n\}_1^\infty\}$$

is called the *lim sup of a , or $(a_t)_{t \in R}$ in density over $\{F_n\}_1^\infty$* .

- (3) We can similarly define $\text{D-lim inf}_{t \in R} a_t$ for any $a : R \rightarrow \mathbb{R}$ by using *ess.g.l.b.*

Clearly, if a , or $(a_t)_{t \in R}$ is bounded, then

$$-\infty < \text{D-lim inf}_{t \in R} a_t \leq \text{D-lim sup}_{t \in R} a_t < \infty$$

over any $\{F_n\}_1^\infty$. In addition, whenever $a \geq 0$ and $\text{D-lim sup}_{t \in R} a_t = 0$, then $\text{D-lim}_{t \in R} a_t = 0$ over $\{F_n\}_1^\infty$. Note that $\text{D-lim}_{t \in R} a_t$ is completely different with the Moore-Smith limit involving a directed system. See [10, Sections 4.2 and 7.2] for the special case of $R = \mathbb{Z}$ or \mathbb{Z}_+ with the discrete topology.

We shall need the following notions and lemma.

Notions and Lemma 1.2. Let \mathcal{F}_+ be the family of all the Borel subsets of R with positive upper density over the same weak Følner sequence $\{F_n\}_1^\infty$ in $(R, +)$. Then

- \mathcal{F}_+ has the “Ramsey property” (cf. [10, Def. 9.1] for $R = \mathbb{Z}$), i.e., if $S_1 \cup S_2 \in \mathcal{F}_+$ and $S_1, S_2 \in \mathcal{B}_R$ then either $S_1 \in \mathcal{F}_+$ or $S_2 \in \mathcal{F}_+$;
- \mathcal{F}_+ has the “van der Corput property”, i.e., for $S \in \mathcal{F}_+$, $\{q \in R \mid (S - q) \cap S \in \mathcal{F}_+\} \in \mathcal{F}_+$.

We define the dual family \mathcal{F}_+^* to consist of all those Borel subsets of R that intersect non-voidly each member of \mathcal{F}_+ . Then $\mathcal{F}_+^* = \{S \in \mathcal{B}_R \mid D(S) = 1 \text{ over } \{F_n\}_1^\infty\}$.

Proof. The Ramsey property of \mathcal{F}_+ trivially holds and the van der Corput property follows clearly from the translation invariance of upper/lower densities over any given weak Følner sequence $\{F_n\}_1^\infty$ in $(R, +)$. \square

We shall need the following lemma, which is a generalization of [10, Lemma 4.8] for the special case that $R = \mathbb{Z}$.

Lemma 1.3. *Let $Q \subset R$ be a Borel subset of density 1 over $\{F_n\}_1^\infty$ and for each $q \in Q$ let R_q be a Borel subset of R of density 1 over $\{F_n\}_1^\infty$. Let S be a Borel subset of R with $D^*(S) > 0$ over $\{F_n\}_1^\infty$ and let $k \geq 1$ be any given integer. Then there exist k distinct points r_1, \dots, r_k in S such that each $r_j - r_i \in Q$ and $r_i \in R_{r_j - r_i}$ for $1 \leq i < j \leq k$.*

Proof. As in the \mathbb{Z} case [10, Lemma 4.8], we shall show by induction that there exists a Borel set $S_k \subset S$ of positive upper density over $\{F_n\}_1^\infty$, and a set of k points t_1, \dots, t_k in R such that for each $r \in S_k$ the k -tuple $r_1 = r + t_1, \dots, r_k = r + t_k$ has the properties sought in the lemma.

For $k = 1$ there is nothing needed to prove by letting $S_1 = S, t_1 = o$; so suppose that S_k and $\{t_1, \dots, t_k\}$ have been found. We shall find t_{k+1} such that $t_{k+1} - t_i \in Q$ for $1 \leq i \leq k$, and a Borel subset S_{k+1} of S_k with positive upper density over $\{F_n\}_1^\infty$ such that for each $r \in S_{k+1}$ the additional conditions, $r + t_{k+1} \in S$ and $r + t_i \in R_{t_{k+1} - t_i}$, $1 \leq i \leq k$, are satisfied.

We first note that since S_k has positive upper density over $\{F_n\}_1^\infty$, then by Lemma 1.2

$$S_k^* = \{t \in R : D^*((S_k - t) \cap S_k) > 0 \text{ over } \{F_n\}_1^\infty\}$$

is of positive upper density over $\{F_n\}_1^\infty$. We now choose t_{k+1} in $S_k^* \cap \bigcap_{i \leq k} (Q + t_i)$. We can then obtain $t_{k+1} - t_i \in Q$ for $1 \leq i \leq k$. Finally we set

$$S_{k+1} = (S_k - t_{k+1}) \cap S_k \cap \bigcap_{i \leq k} (R_{t_{k+1} - t_i} - t_i).$$

This set has positive upper density over $\{F_n\}_1^\infty$ since $(S_k - t_{k+1}) \cap S_k$ does so and $\bigcap_{i \leq k} (R_{t_{k+1} - t_i} - t_i)$ has density 1 over $\{F_n\}_1^\infty$.

The proof of Lemma 1.3 is thus completed. \square

Motivated by [10, Lemmas 4.9, 7.5 and 9.24], [1, Theorem 4.6] and [7, Theorem 7.11], by Lemma 1.3 we can obtain the following useful technical result, which is another version of the classical van der Corput lemma [4].

Lemma 1.4. *Let $x_\cdot : R \rightarrow \mathbf{H}$ be a bounded Borel measurable function, with $\|x_t\| \leq \beta \forall t \in R$, from an lcsc abelian $(R, +)$ into a Hilbert space \mathbf{H} with the inner product $\langle \cdot, \cdot \rangle$ and suppose that*

$$D\text{-}\lim_{r \in R} (D\text{-}\limsup_{t \in R} |\langle x_{r+t}, x_t \rangle|) = 0$$

over a Følner sequence $\{F_n\}_1^\infty$ in $(R, +)$. Then

$$D\text{-}\lim_{t \in R} x_t = 0$$

in \mathbf{H} under the weak topology (namely, $D\text{-}\lim_{t \in R} \langle x_t, x \rangle = 0 \forall x \in \mathbf{H}$) over $\{F_n\}_1^\infty$.

Proof. Without loss of generality, let \mathbf{H} is a real Hilbert space. In contrast to the statement of Lemma 1.4, we can find some number ε with $0 < \varepsilon < 1$ and some unit vector $v \in \mathbf{H}$ such that

$$S := \{t \in R : \langle x_t, v \rangle \geq \varepsilon\} \in \mathcal{F}_+ \text{ over } \{F_n\}_1^\infty.$$

We then for $0 < \delta < \varepsilon^2$ let

$$Q = \left\{ q \in R : D\text{-}\limsup_{t \in R} |\langle x_{q+t}, x_t \rangle| < \frac{\delta}{2} \right\}.$$

So Q has density 1 over $\{F_n\}_1^\infty$ by Def. 1.1 and for each $q \in Q$, the Borel set

$$R_q = \{r \in R: |\langle x_{q+r}, x_r \rangle| < \delta\}$$

has density 1 over $\{F_n\}_1^\infty$. Applying Lemma 1.3 to these sets with k to be specified later. If r_1, \dots, r_k satisfy the conclusion of Lemma 1.3, then we shall have that

$$\langle x_{r_i}, v \rangle > \varepsilon, \quad 1 \leq i \leq k, \quad \text{and} \quad |\langle x_{r_i}, x_{r_j} \rangle| < \delta, \quad 1 \leq i < j \leq k.$$

Set $y_i = x_{r_i} - \varepsilon v$ for $1 \leq i \leq k$. Then

$$\langle y_i, y_j \rangle < \delta - 2\varepsilon + \varepsilon^2 < \delta - \varepsilon^2 < 0, \quad 1 \leq i < j \leq k.$$

But since the y_i are bounded independently of k , and

$$\begin{aligned} 0 &\leq \|y_1 + \dots + y_k\|^2 = \sum_{i=1}^k \|y_i\|^2 + 2 \sum_{i < j} \langle y_i, y_j \rangle \\ &\leq k \max_{1 \leq i \leq k} \|y_i\|^2 - k(k-1)(\varepsilon^2 - \delta) \end{aligned}$$

which tends to $-\infty$ as $k \rightarrow \infty$, we thus arrive at a contradiction.

Therefore the set S in question must have density 0 over $\{F_n\}_1^\infty$ for any x and any $\varepsilon > 0$. This proves Lemma 1.4. \square

The above proof is different with other versions [3, p. 445] and [7, Theorem 7.11]. In addition, let $\{\xi_n; n = 1, 2, \dots\}$ be a sequence of independent random variable on a probability space (Ω, \mathcal{F}, P) with $\xi_n \sim N(0, 1)$ for all $n \geq 1$. Then $\langle \xi_{n+m}, \xi_n \rangle = E_P(\xi_{n+m}\xi_n) = 0$ for $m \neq 0$ and so this shows that in Lemma 1.4, we cannot expect that $D\text{-}\lim_{t \in R} \|x_t\| = 0$.

2. Inverse limits of Kh -systems

In this section, let G be lcsc R -module. Let (X, \mathcal{X}, μ) be any Borel G -space and we write $X = (X, \mathcal{X}, \mu, G)$. For any $T \in G$, we write $tT = T^t$ for any $t \in R$. We will consider a kind of multiple dynamics— Kh -systems.

Motivated by Khintchine's recurrence theorem and [2], we now introduce the following notation:

Definition 2.1. For $X = (X, \mathcal{X}, \mu, G)$, we shall say that X is a Kh -system (Kh is for Khintchine), provided that for any integer $l \geq 2$,

- if $f \in \mathcal{L}^\infty(X, \mathcal{X}, \mu)$ with $f \geq 0$ a.e. and $\int_X f d\mu > 0$, then for any $T_1, \dots, T_l \in G$ and any $\varepsilon > 0$, the set

$$\left\{ t \in R \mid \int_X T_1^t f \cdots T_l^t f d\mu > \left(\int_X f d\mu \right)^l - \varepsilon \right\}$$

is of positive lower Banach density; that is, it is of positive lower density over any weak Følner sequence $\{F_n\}_1^\infty$ in $(R, +)$.

We can then obtain the following result.

Proposition 2.2. *Let $\{\mathcal{X}_\theta; \theta \in \Theta\}$ be a totally ordered family of σ -subalgebras of \mathcal{X} and set $\mathcal{Y} = \sigma(\bigcup_\theta \mathcal{X}_\theta)$. If each $(X, \mathcal{X}_\theta, \mu, G)$ is a Kh-system, then (X, \mathcal{Y}, μ, G) is also a Kh-system.*

Proof. Let $f \in \mathcal{L}^\infty(X, \mathcal{Y}, \mu)$ with $f \geq 0$ a.e. and $\int_X f d\mu > 0$, and let $T_1, \dots, T_l \in G$ and $\varepsilon > 0$. We may assume $0 \leq f(x) \leq 1/2$ a.e. without loss of generality. Let $\epsilon > 0$ be any given with $\epsilon < \varepsilon/2$. Since $\mathcal{Y} = \sigma(\bigcup_\theta \mathcal{X}_\theta)$ and $\{\mathcal{X}_\theta\}$ is totally ordered, then it follows that one can find some $\tilde{f} \in \mathcal{L}^\infty(X, \mathcal{X}_\theta, \mu)$ with $\|f - \tilde{f}\|_2 \ll \epsilon$ for some $\theta \in \Theta$. Simply write $\tilde{f} = E_\mu(f|\mathcal{X}_\theta)$. Then $0 \leq \tilde{f} \leq 1/2$ a.e. and $\|f - \tilde{f}\|_2 \ll \epsilon$. Hence we can set $f = \tilde{f} + \phi_\theta$ where $|\phi_\theta| \leq 1$ a.e. and $\|\phi_\theta\|_2 \ll \epsilon$, and further for any $t \in R$

$$T_1^t f \cdots T_l^t f = T_1^t \tilde{f} \cdots T_l^t \tilde{f} + \dots, \quad \|\dots\|_2 < \epsilon$$

and further for any $t \in R$

$$\int_X T_1^t f \cdots T_l^t f d\mu \geq \int_X T_1^t \tilde{f} \cdots T_l^t \tilde{f} d\mu - \epsilon.$$

Since $(X, \mathcal{X}_\theta, \mu, G)$ is a Kh-system and $\int_X f d\mu = \int_X \tilde{f} d\mu > 0$, then the set

$$\left\{ t \in R \mid \int_X T_1^t \tilde{f} \cdots T_l^t \tilde{f} d\mu > \left(\int_X \tilde{f} d\mu \right)^l - \epsilon \right\}$$

is of positive lower density over any weak Følner sequence $\{F_n\}_1^\infty$ in $(R, +)$. This implies the statement by Def. 2.1 for $\epsilon < \varepsilon/2$.

The proof of Proposition 2.2 is therefore completed. \square

This result immediately leads to the following important fact.

Corollary 2.3. *Let $\{\mathbf{X}_\xi\}$ be a Furstenberg factors chain of $\mathbf{X} = (X, \mathcal{X}, \mu, G)$ in accordance with Theorem 0.2. Assume that the ordinal ξ is a limit ordinal and that \mathbf{X}_θ is a Kh-system for each $\theta < \xi$. Then \mathbf{X}_ξ is a Kh-system.*

Proof. According to Theorem 0.2, $\{\mathcal{X}_\theta; \theta < \xi\}$ is totally ordered. Then the statement follows from the foregoing Proposition 2.2. \square

Proposition 2.2 shows that the Kh-property survives taking inverse limits.

3. Kh-property of totally relatively weak-mixing extensions

Let G be an lcsc R -module with the identity e and with the continuous scalar multiplication $(t, g) \mapsto g^t$ of $R \times G$ to G . Let there be any given a short factors series:

$$\mathbf{X} = (X, \mathcal{X}, \mu, G) \xrightarrow{Id_X} \mathbf{X}' = (X, \mathcal{X}', \mu, G) \xrightarrow{\pi} \mathbf{Y} = (Y, \mathcal{Y}, \nu, G),$$

where (X, \mathcal{X}, μ) is a standard Borel G -space so we can decompose $\mu = \int_Y \mu_y d\nu(y)$ and where $\mathbf{Y} = (Y, \mathcal{Y}, \nu, G)$. Based on the measure-preserving system $\mathbf{X} = (X, \mathcal{X}, \mu, G)$, we regard every element g of G as a μ -preserving transformation of X onto itself and we also often identify g with the unitary operator $U_g: \phi \mapsto \phi \circ g$ of the \mathcal{L}^p -function spaces if no confusion, for $p \geq 1$.

3.1. Relatively weak-mixing extensions

Recall from the viewpoint of group as in [11, 10] that an extension $\pi: X \rightarrow Y$ is referred to as *relatively weak-mixing for an element g in G* if every g -invariant (or equivalently $\{g^n \mid n \in \mathbb{Z}\}$ -invariant) function H in $\mathcal{L}^2(X, \mathcal{X}, \mu) \otimes_Y \mathcal{L}^2(X, \mathcal{X}, \mu)$ is a function on (Y, \mathcal{Y}, ν) via the relative-product factor G -map

$$\pi \times_Y \pi: X \times X \rightarrow Y; \quad (x_1, x_2) \mapsto x_1.$$

Then we say $\pi: X \rightarrow Y$ is *relatively weak-mixing for G* if $\pi: X \rightarrow Y$ is relatively weak-mixing for every $g, g \neq e$, in G (cf. [10, Def. 6.3]).

Now for any subset S of the R -module G , by $\langle S \rangle_R$ we denote the submodule generated by the elements $g \in S$ over the ring $(R, +, \cdot)$. We note that for any $H \in \mathcal{L}^2(X, \mathcal{X}, \mu) \otimes_Y \mathcal{L}^2(X, \mathcal{X}, \mu)$, it is g -invariant if and only if it is $\langle g \rangle_{\mathbb{Z}}$ -invariant. However, this is not the case about $\langle g \rangle_R$ -invariance when $g \neq e$ in general. More generally than the above relatively weak-mixing in the literature, we now will introduce other relatively weak-mixing condition for R -modules.

Definition 3.1 ([6]). Let Γ be a subset of the *lcsc* R -module G with $\Gamma \neq \{e\}$ and $g \in G, g \neq e$. Then $\pi: X' \rightarrow Y$ is said to be

- (a) *relatively weak-mixing for g* if every $\langle g \rangle_R$ -invariant $H \in \mathcal{L}^2(X, \mathcal{X}', \mu) \otimes_Y \mathcal{L}^2(X, \mathcal{X}', \mu)$ is a function on (Y, \mathcal{Y}, ν) via $\pi \times_Y \pi: X \times X \rightarrow Y$.
- (b) *totally relatively weak-mixing for Γ* if $\pi: X' \rightarrow Y$ is relatively weak-mixing for each g in Γ with $g \neq e$ under the sense of (a);
- (c) *jointly relatively ergodic for Γ* if every $\langle \Gamma \rangle_R$ -invariant function $\varphi \in \mathcal{L}^2(X, \mathcal{X}', \mu)$ is a function on (Y, \mathcal{Y}, ν) via $\pi: X' \rightarrow Y$.
- (d) *jointly relatively weak-mixing for Γ* if $\pi \times_Y \pi: X' \times_Y X' \rightarrow Y$ is jointly relatively ergodic for Γ in the sense of (c).

Since every $\langle \Gamma \rangle_R$ -invariant function is $\langle g \rangle_R$ -invariant for each $g \in \Gamma$, a totally relatively weak-mixing extension must be a jointly relatively weak-mixing extension for Γ ; but the converse does not need to be true by considering a one-point factor G -system Y .

In addition, if $\pi: X' \rightarrow Y$ is relatively weak-mixing for g , then it is not necessarily totally relatively weak-mixing for $\langle g \rangle_R$ unless $rR = R$ for each $r \neq 0$ like R to be a field.

When $(R, +, \cdot) = (\mathbb{Z}, +, \cdot)$, Def. 3.1-(a) and (b) coincide exactly with the classical case in the literature [11, 10]. However, if $(R, +, \cdot) \neq (\mathbb{Z}, +, \cdot)$, then our relatively weak-mixing is weaker than that of [11, 10].

Lemma 3.2. *If $\pi: X' = (X, \mathcal{X}', \mu, G) \rightarrow Y = (Y, \mathcal{Y}, \nu, G)$ is a relatively ergodic extension for an element T in G , then for any $\varphi \in \mathcal{L}^2(X, \mathcal{X}', \mu)$ with $E_\mu(\varphi|\pi^{-1}[\mathcal{Y}]) = 0$ (or equivalently $E_\mu(\varphi|Y) = 0$), we have*

$$\lim_{n \rightarrow \infty} \frac{1}{|F_n|} \int_{F_n} T^t \varphi dt = 0$$

in \mathcal{L}^2 -norm, over any weak Følner sequence $\{F_n\}_1^\infty$ in $(R, +)$.

Proof. Let $\mathcal{X}'_{\mu, \langle T \rangle_R} = \{A \in \mathcal{X}' \mid \mu(T^{-t}(A) \triangle A) = 0 \ \forall t \in R\}$. By the \mathcal{L}^2 -mean ergodic theorem (cf. [7] or more precisely [5]), it follows that

$$\mathcal{L}^2(\mu)\text{-}\lim_{n \rightarrow \infty} \frac{1}{|F_n|} \int_{F_n} T^t \varphi dt = E_\mu(\varphi | \mathcal{X}'_{\mu, \langle T \rangle_R}).$$

By hypothesis $\pi: X' \rightarrow Y$ is relatively ergodic for the element T so that $\mathcal{X}'_{\mu, \langle T \rangle_R} \subseteq \pi^{-1}[\mathcal{Y}]$ and then $E_\mu(\varphi|\mathcal{X}'_{\mu, \langle T \rangle_R}) = E_\mu(E_\mu(\varphi|\pi^{-1}[\mathcal{Y}]|\mathcal{X}'_{\mu, \langle T \rangle_R})) \equiv 0$ a.e. which yields the desired result. \square

This lemma is a generalization of [10, Lemma 6.1]. The following consequence of it is a generalization of [11, Lemma 1.3] or [10, Proposition 6.2] (also see [7, Proposition 7.30]).

Proposition 3.3. *If $\pi: X' \rightarrow Y$ is relatively weak-mixing for some T , where $T \in G$ with $T \neq e$, and $\phi, \psi \in \mathcal{L}^\infty(X, \mathcal{X}', \mu)$, then*

$$\lim_{n \rightarrow \infty} \frac{1}{|F_n|} \int_{F_n} \int_Y |E_\mu(\psi T^t \phi | Y)(y) - E_\mu(\psi | Y) T^t E_\mu(\phi | Y)(y)|^2 \nu(dy) dt = 0,$$

over any weak Følner sequence $\{F_n\}_1^\infty$ in $(R, +)$.

Proof. Take $\phi, \psi \in \mathcal{L}^\infty(X, \mathcal{X}', \mu)$ so that $\phi \otimes \bar{\phi}, \psi \otimes \bar{\psi} \in \mathcal{L}^2(X \times X, \mathcal{X}' \otimes_Y \mathcal{X}', \mu \otimes_Y \mu)$. By [10, Proposition 5.12],

$$\begin{aligned} E_{\mu \otimes_Y \mu}((\psi \otimes \bar{\psi}) T^t (\phi \otimes \bar{\phi}) | Y) &= E_{\mu \otimes_Y \mu}((\psi T^t \phi) \otimes (\bar{\psi} T^t \bar{\phi}) | Y) \\ &= |E_\mu(\psi T^t \phi | Y)|^2. \end{aligned}$$

First assume that $E_\mu(\phi | Y) = 0$, then $E_{\mu \otimes_Y \mu}(\phi \otimes \bar{\phi} | Y) = |E_\mu(\phi | Y)|^2 = 0$. Then by Lemma 3.2 and the above equality, we can obtain

$$\lim_{n \rightarrow \infty} \frac{1}{|F_n|} \int_{F_n} \int_Y |E_\mu(\psi T^t \phi | Y)(y)|^2 \nu(dy) dt = 0.$$

Now let ϕ be an arbitrary bounded function. Then the function $\hat{\phi} = \phi - E_\mu(\phi|\pi^{-1}[\mathcal{Y}])$ satisfies $E_\mu(\hat{\phi} | Y) = 0$. Replacing ϕ by $\hat{\phi}$ in the above equality leads to what we need from that

$$E_\mu(\psi T^t E_\mu(\phi|\pi^{-1}[\mathcal{Y}]) | Y) = E_\mu(\psi | Y) T^t E_\mu(\phi | Y).$$

This proves Proposition 3.3. \square

The following is a generalization of [11, Proposition 1.2] or [10, Proposition 6.3].

Lemma 3.4. *If $\pi: X' \rightarrow Y$ is relatively weak-mixing for some T , where $T \in G, T \neq e$, and $\pi_2: X'_2 \rightarrow Y$ is relatively ergodic for T , then $\pi \times_Y \pi_2: X' \times_Y X'_2 \rightarrow Y$ is relatively ergodic for T . Particularly, $\pi \times_Y \pi: X' \times_Y X' \rightarrow Y$ is relatively weak-mixing for T .*

Proof. The statement follows from a slight modification of the proof of [11, Proposition 1.2]. So we omit the details here. \square

The above arguments are still valid for the jointly relatively weak-mixing case. Specially we have the following

Proposition 3.5. *If $\pi: X' \rightarrow Y$ is jointly relatively weak-mixing for G itself, and $\phi, \psi \in \mathcal{L}^\infty(X, \mathcal{X}', \mu)$, then*

$$\lim_{n \rightarrow \infty} \frac{1}{m_G(F_n)} \int_{F_n} \int_Y |E_\mu(\psi g \phi | Y) - E_\mu(\psi | Y) g E_\mu(\phi | Y)|^2 dv dg = 0,$$

over any weak Følner sequence $\{F_n\}_1^\infty$ in G , where m_G and dg are the same left Haar measure of the lsc abelian group G .

3.2. *Kh-property of totally relatively weak-mixing extensions*

We will now consider the lifting of *Kh*-property of totally relatively weak-mixing extensions. Following Def. 3.1 and Def. 1.1 we begin with a lemma, which generalizes [10, Lemmas 7.3 and 7.6 for $R = \mathbb{Z}$]. Also see [7, Proposition 7.30] for $G = R = \mathbb{Z}$.

Lemma 3.6. *If $\pi: X' = (X, \mathcal{X}', \mu, G) \rightarrow Y = (Y, \mathcal{Y}, \nu, G)$ is relatively weak-mixing for some T where $T \in G$ with $T \neq e$, and if $\psi \in \mathcal{L}^2(X, \mathcal{X}', \mu)$ and ϕ belongs to $\mathcal{L}^\infty(X, \mathcal{X}', \mu)$ with $E_\mu(\phi|\pi^{-1}[\mathcal{Y}]) = 0$ or equivalently $E_\mu(\phi|Y) = 0$, then*

$$\text{D-lim}_{t \in R} \int_X \psi T^t \phi d\mu = 0 \quad \text{and} \quad \text{D-lim}_{t \in R} \|E_\mu(\psi T^t \phi|Y)\|_{2,\nu} = 0$$

over any weak Følner sequence $\{F_n\}_1^\infty$ in $(R, +)$.

Proof. Over any weak Følner sequence $\{F_n\}_1^\infty$ in $(R, +)$, according to Proposition 3.3 we have, for any $\psi \in \mathcal{L}^\infty(X, \mathcal{X}', \mu)$,

$$\lim_{n \rightarrow \infty} \frac{1}{|F_n|} \int_{F_n} \int_Y |E_\mu(\psi T^t \phi|Y)(y)|^2 d\nu(y) dt = 0$$

so that

$$0 = \lim_{n \rightarrow \infty} \frac{1}{|F_n|} \int_{F_n} \left| \int_Y E_\mu(\psi T^t \phi|Y) d\nu \right|^2 dt = \lim_{n \rightarrow \infty} \frac{1}{|F_n|} \int_{F_n} \left| \int_X \psi T^t \phi d\mu \right|^2 dt.$$

This implies that $\text{D-lim}_{t \in R} \int_X \psi T^t \phi d\mu = 0$ and $\text{D-lim}_{t \in R} \|E_\mu(\psi T^t \phi|Y)\|_{2,\nu} = 0$ over $\{F_n\}_1^\infty$ for a dense set of ψ .

Thus the proof of Lemma 3.6 is completed. \square

The following result is important for us to prove that a totally relatively weak-mixing extension of a *Kh*-system is also a *Kh*-system, which is a generalization of [10, Proposition 7.4] (also see [7, Theorem 7.27 and Proposition 7.31] for the special case that $G = R = \mathbb{Z}$).

Proposition 3.7. *Let $\pi: X' \rightarrow Y$ be totally relatively weak-mixing for G and let $T_1, \dots, T_l \in G$ be distinct with $T_i \neq e$ for $1 \leq i \leq l$. If $f_1, \dots, f_l \in \mathcal{L}^\infty(X, \mathcal{X}', \mu)$, then in $\mathcal{L}^2(X, \mathcal{X}', \mu)$ with the **weak** topology, it holds that*

$$\text{D-lim}_{t \in R} \left\{ \prod_{i=1}^l T_i^t f_i - \prod_{i=1}^l T_i^t E_\mu(f_i|\pi^{-1}[\mathcal{Y}]) \right\} = 0$$

over any weak Følner sequence in $(R, +)$.

Proof. Given any distinct elements $T_1, \dots, T_l \in G$ with $T_i \neq e$ for $1 \leq i \leq l$ and let $\{F_n\}_1^\infty$ be an arbitrary weak Følner sequence in $(R, +)$. We will prove by induction on l that over this weak Følner sequence,

$$\prod_{i=1}^l T_i^t f_i - \prod_{i=1}^l T_i^t E_\mu(f_i|\pi^{-1}[\mathcal{Y}]) \xrightarrow{\text{weakly, D-}} 0 \quad (3.1)$$

for any $f_1, \dots, f_l \in \mathcal{L}^\infty(X, \mathcal{X}', \mu)$.

We first consider the case of $l = 1$. Let $\phi = f_1 - E_\mu(f_1|\pi^{-1}[\mathcal{Y}])$ and set $\phi_t = T_1^t \phi$. Then by Lemma 3.6, it follows that

$$\text{D-}\lim_{r \in R} \left(\text{D-}\lim_{t \in R} |\langle \phi_{r+t}, \phi_t \rangle| \right) = \text{D-}\lim_{r \in R} \left(\text{D-}\lim_{t \in R} \left| \int_X \phi T_1^r \phi d\mu \right| \right) = 0.$$

Thus $\text{D-}\lim_{t \in R} \{T_1^t f_1 - T_1^t E_\mu(f_1|\pi^{-1}[\mathcal{Y}])\} = \text{D-}\lim_{t \in R} \phi_t = 0$ weakly by Lemma 1.4.

Next, we assume T_1, \dots, T_l are distinct elements in G with $T_i \neq e$ for $1 \leq i \leq l$ and $l \geq 2$. By applying to the foregoing expression (3.1) the telescoping sum identity

$$\prod_{i=1}^l a_i - \prod_{i=1}^l b_i = \sum_{j=1}^l \left(\prod_{i=1}^{j-1} a_i \right) (a_j - b_j) \left(\prod_{i=j+1}^l b_i \right) \quad (3.2)$$

where and in the sequel an empty product is always interpreted as 1, we reduce our proof of the above (3.1) to the proof of

$$\text{weakly D-}\lim_{t \in R} \prod_{i=1}^l T_i^t f_i = 0, \quad \forall f_1, \dots, f_l \in \mathcal{L}^\infty(X, \mathcal{X}', \mu)$$

under the additional hypothesis that for some $1 \leq j \leq l$, $E_\mu(f_j|\pi^{-1}[\mathcal{Y}]) = 0$ since we have

$$E_\mu(f_j - E_\mu(f_j|\pi^{-1}[\mathcal{Y}])) = 0.$$

Without loss of generality assume $E_\mu(f_l|\pi^{-1}[\mathcal{Y}]) = 0$, and the van der Corput-type lemma (Lemma 1.4) may be employed as follows. Set

$$\psi_t = \prod_{i=1}^l T_i^t f_i \in \mathcal{L}^2(X, \mathcal{X}', \mu) \quad \forall t \in R.$$

Then for any $r, t \in R$, we can get that

$$\langle \psi_{r+t}, \psi_t \rangle = \int_X \prod_{i=1}^l T_i^{r+t} f_i \prod_{i=1}^l T_i^t f_i d\mu = \int_X (f_l T_l^r f_l) \prod_{i=1}^{l-1} (T_i T_l^{-1})^t (f_i T_i^r f_i) d\mu.$$

Since T_1, T_2, \dots, T_l are distinct, we can apply the induction hypothesis to obtain

$$\text{D-}\lim_{t \in R} \left\{ \langle \psi_{r+t}, \psi_t \rangle - \int_X (f_l T_l^r f_l) \prod_{i=1}^{l-1} (T_i T_l^{-1})^t E_\mu(f_i T_i^r f_i |\pi^{-1}[\mathcal{Y}]) d\mu \right\} = 0$$

so that

$$\text{D-}\limsup_{t \in R} |\langle \psi_{r+t}, \psi_t \rangle| = \text{D-}\limsup_{t \in R} \left| \int_X (f_l T_l^r f_l) \prod_{i=1}^{l-1} (T_i T_l^{-1})^t E_\mu(f_i T_i^r f_i |\pi^{-1}[\mathcal{Y}]) d\mu \right|.$$

Then by the self-adjointness of conditional expectation (cf. [12, Theorem 6.1(vi)]), we have

$$\begin{aligned} & \int_X (f_l T_l^r f_l) \prod_{i=1}^{l-1} (T_i T_l^{-1})^t E_\mu(f_i T_i^r f_i |\pi^{-1}[\mathcal{Y}]) d\mu \\ &= \int_X E_\mu(f_l T_l^r f_l |\pi^{-1}[\mathcal{Y}]) \prod_{i=1}^{l-1} (T_i T_l^{-1})^t E_\mu(f_i T_i^r f_i |\pi^{-1}[\mathcal{Y}]) d\mu. \end{aligned}$$

Since $\|\prod_{i=1}^{l-1} T_i^t T_l^{-t} (f_i T_i^r f_i)\|_\infty \leq C$ for some constant $C > 0$, then by the Hölder inequality it follows that

$$\left| \int_X (f_l T_l^r f_l) \prod_{i=1}^{l-1} (T_i T_l^{-1})^t E_\mu(f_i T_i^r f_i | \pi^{-1}[\mathcal{Y}]) d\mu \right| \leq C \|E_\mu(f_l T_l^r f_l | \mathbf{Y})\|_{2,v}$$

so that

$$\text{D-}\limsup_{t \in R} |\langle \psi_{r+t}, \psi_t \rangle| \leq C \|E_\mu(f_l T_l^r f_l | \mathbf{Y})\|_{2,v} \quad \forall r \in R$$

and further by Lemma 3.6

$$\text{D-}\lim_{r \in R} \left(\text{D-}\limsup_{t \in R} |\langle \psi_{r+t}, \psi_t \rangle| \right) \leq \text{D-}\lim_{r \in R} C \|E_\mu(f_l T_l^r f_l | \mathbf{Y})\|_{2,v} = 0.$$

Hence Lemma 1.4 follows that $\text{D-}\lim_{t \in R} \psi_t = 0$ over $\{F_n\}_1^\infty$ in the sense of the weak topology of $\mathfrak{L}^2(X, \mathcal{X}', \mu)$.

The proof of Proposition 3.7 is thus completed. \square

The following result is a generalization of the classical \mathbb{Z} -module case (cf. [9, Theorem 2.3], [10, Theorem 4.10] and [1, Theorem 4.5] by using different induction approaches).

Corollary 3.8. *Let $X' = (X, \mathcal{X}', \mu, G)$ be totally weak-mixing. Then for any distinct elements T_1, \dots, T_l in G with $T_i \neq e$ for $1 \leq i \leq l$ and any f_1, \dots, f_l in $\mathfrak{L}^\infty(X, \mathcal{X}', \mu)$ we have*

$$\text{D-}\lim_{t \in R} \prod_{i=1}^l T_i^t f_i = \left(\int_X f_1 d\mu \right) \cdots \left(\int_X f_l d\mu \right) \quad \text{weakly in } \mathfrak{L}^2(X, \mathcal{X}, \mu)$$

over any weak Følner sequence in $(R, +)$.

Proof. Since X is a totally relatively weak-mixing extension of the one-point system for G , hence the statement follows from Proposition 3.7. \square

As a consequence of Proposition 3.7 we have the following, which is a generalization of [10, Proposition 7.7].

Lemma 3.9. *If $\pi: X' \rightarrow Y$ is totally relatively weak-mixing for G and $T_1, \dots, T_l \in G$ are distinct with $T_i \neq e$ for $1 \leq i \leq l$, then for any $f_0 \in \mathfrak{L}^2(X, \mathcal{X}', \mu)$ and $f_1, \dots, f_l \in \mathfrak{L}^\infty(X, \mathcal{X}', \mu)$,*

$$\text{D-}\lim_{t \in R} \left\{ \int_X f_0 \prod_{i=1}^l T_i^t f_i d\mu - \int_Y E_\mu(f_0 | \mathbf{Y}) \prod_{i=1}^l T_i^t E_\mu(f_i | \mathbf{Y}) d\nu \right\} = 0,$$

over any weak Følner sequence in $(R, +)$.

Proof. Over any weak Følner sequence in $(R, +)$, we rewrite this as

$$\text{D-}\lim_{t \in R} \left\{ \int_X f_0 \prod_{i=1}^l T_i^t f_i d\mu - \int_X E_\mu(f_0 | \pi^{-1}[\mathcal{Y}]) \prod_{i=1}^l T_i^t E_\mu(f_i | \pi^{-1}[\mathcal{Y}]) d\mu \right\} = 0$$

and noting that by the self-adjointness of the conditional expectation operator $E_\mu(\cdot|\pi^{-1}[\mathcal{Y}])$

$$\int_X E_\mu(f_0|\pi^{-1}[\mathcal{Y}]) \prod_{i=1}^l T_i^t E_\mu(f_i|\pi^{-1}[\mathcal{Y}]) d\mu = \int_X f_0 \prod_{i=1}^l T_i^t E_\mu(f_i|\pi^{-1}[\mathcal{Y}]) d\mu,$$

whence we only need to prove that

$$\text{D-}\lim_{t \in R} \left\langle f_0, \prod_{i=1}^l T_i^t f_i - \prod_{i=1}^l T_i^t E_\mu(f_i|\pi^{-1}[\mathcal{Y}]) \right\rangle = 0.$$

However, this follows at once from Proposition 3.7. \square

The following is a simple consequence of Lemma 3.9.

Proposition 3.10. *Let $\pi: X' \rightarrow Y$ be totally relatively weak-mixing for G and let Y be a Kh -system; then X' is also a Kh -system.*

Finally we remark that for a jointly relatively weak-mixing extension $\pi: X' \rightarrow Y$ (cf. Def. 3.1-(d)) there exists no such an useful theory in the literature.

4. Kh -property of chaotic G -systems

In this section, let G be an lcsc R -module and write the continuous scalar multiplication as $(t, T) \mapsto T^t$ from $R \times G$ to G . Let $X' = (X, \mathcal{X}', \mu, G)$ be a (not necessarily standard) Borel G -system.

4.1. Compact G -system

Different with to say that X is a compact space or X' is relatively compact, by a *compact G -space* (X, \mathcal{X}', μ) or a *compact G -system* $X' = (X, \mathcal{X}', \mu, G)$, we mean that for each f in $\mathcal{L}^2(X, \mathcal{X}', \mu)$, the G -orbit $G[f] = \{Tf \mid T \in G\}$ is precompact/totally-bounded in the Hilbert space $\mathcal{L}^2(X, \mathcal{X}', \mu)$.

4.2. Chaotic G -system

We shall say $X' = (X, \mathcal{X}', \mu, G)$ is **chaotic** if there exists a direct product $G = G_c \times G_w$ of two nontrivial R -submodules of G such that $X'_c = (X, \mathcal{X}', \mu, G_c)$ is a compact G_c -system and $X'_w = (X, \mathcal{X}', \mu, G_w)$ is a totally weak-mixing G_w -system. This is stronger than a primitive system in the Furstenberg factors chain of X .

The following result will play an important role in proving Theorem 0.1.

Proposition 4.1. *If $X' = (X, \mathcal{X}', \mu, G)$ is chaotic, then it is a Kh -system such that for any f in $\mathcal{L}^\infty(X, \mathcal{X}', \mu)$ with $f \geq 0$ a.e. and $\int_X f d\mu > 0$ and for any $\varepsilon > 0$,*

$$\liminf_{n \rightarrow +\infty} \frac{1}{|F_n|} \int_{F_n} \left(\int_X T_1^t f \cdots T_l^t f d\mu \right) dt > \left(\int_X f d\mu \right)^l - \varepsilon \quad (4.1)$$

over any weak Følner sequence $\{F_n\}_1^\infty$ in $(R, +)$, for any $T_1, \dots, T_l \in G$, $l \geq 2$.

Proof. Let $f \in \mathfrak{L}^\infty(X, \mathcal{X}', \mu)$ be such that $f \geq 0$ a.e. and $\int_X f d\mu > 0$. Given any weak Følner sequence $\{F_n\}_1^\infty$ in $(R, +)$ and any $T_1, \dots, T_l \in G$ where $l \geq 2$. Since $G = G_c \times G_w$, we may write

$$T_i = S_i H_i, \quad S_i \in G_w, \quad H_i \in G_c, \quad i = 1, \dots, l.$$

Moreover, there is no loss of generality in assuming

$$0 \leq f(x) \leq \frac{1}{2} \quad (\mu\text{-a.e.}).$$

As $\mathbf{X}'_w = (X, \mathcal{X}', \mu, G_w)$ is totally weak-mixing, it follows from Corollary 3.8 and Jensen's inequality that

$$\text{D-}\lim_{t \in R} \int_X S_1^t f \cdots S_l^t f d\mu > \left(\int_X f d\mu \right)^l - \frac{\varepsilon}{2}.$$

Thus, the set

$$\mathcal{S} = \left\{ t \in R \mid \int_X S_1^t f \cdots S_l^t f d\mu > \left(\int_X f d\mu \right)^l - \frac{\varepsilon}{2} \right\}$$

is of density 1 over $\{F_n\}_1^\infty$.

On the other hand, since $\mathbf{X}'_c = (X, \mathcal{X}', \mu, G_c)$ is a compact G_c -system, it follows that for arbitrary $\varepsilon > 0$,

$$\mathcal{H} = \left\{ t \in R \mid \max_{1 \leq i \leq l} \|H_i^t f - f\|_2 < \frac{\varepsilon}{2^{l+1}} \right\}$$

is of positive lower density over $\{F_n\}_1^\infty$. Indeed, since $f(x)$ is almost periodic for G_c , the orbits $G_c[f]$ is precompact in $\mathfrak{L}^2(X, \mathcal{X}', \mu)$. Hence the partial orbits $\langle H_i \rangle_R[f]$, for $1 \leq i \leq l$, all are also precompact in $\mathfrak{L}^2(X, \mathcal{X}', \mu)$. Write

$$W = \overbrace{\mathfrak{L}^2(X, \mathcal{X}', \mu) \times \cdots \times \mathfrak{L}^2(X, \mathcal{X}', \mu)}^{l\text{-times}}.$$

By considering the diagonal R -action $\varphi: R \times W \rightarrow W$ given by

$$(t, (\phi_1, \dots, \phi_l)) \mapsto (H_1^t \phi_1, \dots, H_l^t \phi_l),$$

and using Tychonoff's Theorem it follows that the l -tuple partial orbit of the point (f, \dots, f)

$$R_\varphi[(f, \dots, f)] = \{(H_1^t f, \dots, H_l^t f) \in W \mid t \in R\}$$

is precompact in W and thus one can find a finite set of points, say t_1, \dots, t_k , in R such that

$$\min_{1 \leq j \leq k} \max_{1 \leq i \leq l} \|H_i^{t_j} f - H_i^t f\|_2 < \frac{\varepsilon}{2^{l+1}} \quad \forall t \in R$$

so that the Borel set

$$\mathcal{H} = \left\{ t \in R \mid \max_{1 \leq i \leq l} \|H_i^t f - f\|_2 < \frac{\varepsilon}{2^{l+1}} \right\} \quad (4.2)$$

is **strongly syndetic** in $(R, +)$ in the sense that $\mathcal{H} + \{t_1, \dots, t_k\} = R$. This implies that \mathcal{H} has positive lower density over $\{F_n\}_1^\infty$ in $(R, +)$; this is because otherwise, there would be a subsequence of $\{n\}$, say $n_\ell \rightarrow \infty$, such that

$$\begin{aligned} 0 &= \lim_{\ell \rightarrow \infty} \frac{|\mathcal{H} \cap F_{n_\ell}|}{|F_{n_\ell}|} = \lim_{\ell \rightarrow \infty} \frac{|\mathcal{H} \cap (F_{n_\ell} - t_j)|}{|F_{n_\ell}|} = \lim_{\ell \rightarrow \infty} \frac{|(\mathcal{H} + t_j) \cap F_{n_\ell}|}{|F_{n_\ell}|} \quad (1 \leq j \leq k) \\ &= \lim_{\ell \rightarrow \infty} \frac{|(\mathcal{H} + K) \cap F_{n_\ell}|}{|F_{n_\ell}|} \quad (\text{where } K = \{t_1, \dots, t_k\}) \\ &= 1 \end{aligned}$$

which is a contradiction to $\mathcal{H} + \{t_1, \dots, t_k\} = R$.

Then for $1 \leq i \leq l$, we can express $H_i^t f$ as

$$H_i^t f(x) = f(x) + \psi_{i,t}(x) \quad \forall t \in \mathcal{H}, \quad \text{where } |\psi_{i,t}(x)| \leq 1, \quad \|\psi_{i,t}\|_2 < \frac{\varepsilon}{2^{l+1}}.$$

Next we define $P = \mathcal{S} \cap \mathcal{H}$ which is of positive lower density over $\{F_n\}_1^\infty$. Then for any $t \in P$,

$$\begin{aligned} \int_X T_1^t f \cdots T_l^t f \, d\mu &= \int_X S_1^t(H_1^t f) \cdots S_l^t(H_l^t f) \, d\mu \\ &\geq \int_X S_1^t f \cdots S_l^t f \, d\mu - \frac{\varepsilon}{2} \\ &> \left(\int_X f \, d\mu \right)^l - \varepsilon. \end{aligned}$$

This implies (4.1) for $\varepsilon > 0$ is arbitrary.

This proves Proposition 4.1. □

4.3. Proof of Theorem 0.1

Let G be an lcscN R -module over a syndetic ring $(R, +, \cdot)$ and let (X, \mathcal{X}, μ) be a nontrivial standard Borel G -space. Let $\mathbf{X}_1 = (X, \mathcal{X}_1, \mu, G)$ be the first intermediate factor in the Furstenberg factors chain of $\mathbf{X} = (X, \mathcal{X}, \mu, G)$ by Theorem 0.2. Then by Proposition 4.1, it follows that for any $A \in \mathcal{X}_1$ with $\mu(A) > 0$, any $g_1, \dots, g_l \in G$ with any $l \geq 2$ and any $\varepsilon > 0$, the set

$$N_{g_1, \dots, g_l}(A, \varepsilon) = \left\{ t \in R \mid \int_X g_1^t 1_A \cdots g_l^t 1_A \, d\mu > \mu(A)^l - \varepsilon \right\},$$

is of positive lower Banach density in $(R, +)$; i.e.,

$$\liminf_{n \rightarrow +\infty} \frac{|F_n \cap N_{g_1, \dots, g_l}(A, \varepsilon)|}{|F_n|} > 0 \quad (4.3)$$

over any weak Følner sequence $\{F_n\}_1^\infty$ in $(R, +)$.

We now only need to show that $N_{g_1, \dots, g_l}(A, \varepsilon)$ is syndetic in $(R, +)$; that is, there is a compact subset K of R with $N_{g_1, \dots, g_l}(A, \varepsilon) + K = R$. To the contrary, for any weak Følner sequence $\{K_n\}_1^\infty$ in $(R, +)$, there exists a sequence of elements $t_n \in R$ so that $N_{g_1, \dots, g_l}(A, \varepsilon) \cap (t_n + K_n) = \emptyset$ for all $n \geq 1$. Set $F_n = t_n + K_n$ and it is easy to see that $\{F_n\}_1^\infty$ is also a weak Følner sequence in $(R, +)$. But

$$\int_X g_1^t 1_A \cdots g_l^t 1_A \, d\mu \leq \mu(A)^l - \varepsilon \quad \forall t \in F_n, \quad n \geq 1.$$

This is a contradiction to (4.3) and thus the proof of Theorem 0.1 is completed.

It should be noted that if (4.3) holds only over the Følner sequence $\{F_n\}_1^\infty$ satisfying the Tempelman or Shulman conditions, then we cannot deduce the syndetic property of $N_{g_1, \dots, g_l}(A, \varepsilon)$ in $(R, +)$ since $F_n = t_n + K_n$ in the above is not necessarily a Følner sequence satisfying these conditions.

Acknowledgments

This work was partly supported by National Natural Science Foundation of China grant #11271183 and PAPD of Jiangsu Higher Education Institutions.

References

- [1] V. BERGELSON. *Combinatorial and Diophantine Applications of Ergodic Theory*. Handbook of Dynamical Systems, Vol. **1B**. Eds. B. Hasselblatt and A. Katok. Elsevier, Amsterdam, 2006, pp. 745–869.
- [2] V. BERGELSON, B. HOST AND B. KRA. Multiple recurrence and nilsequences. *Invent. Math.*, **160** (2005), 261–303.
- [3] V. BERGELSON AND A. LEIBMAN. A nilpotent Roth theorem. *Invent. Math.*, **147** (2002), 429–470.
- [4] J. G. VAN DER CORPUT. *Diophantische Ungleichungen. I. Zur Gleichverteilung Modulo Eins*. *Acta Math.*, **56** (1931), 373–456.
- [5] X. DAI. Pointwise convergence of ergodic averages of bounded measurable functions for amenable groups. arXiv: 1604.00611v2 [math.DS].
- [6] X. DAI. An extension of Furstenberg’s structure theorem for Noetherian modules and multiple recurrence theorems I. arXiv: 1605.08592v1 [math.DS].
- [7] M. EINSIEDLER AND T. WARD. *Ergodic Theory with a view towards Number Theory*. GTM 259, Springer-Verlag, 2011.
- [8] N. FRANTZIKINAKIS. Multiple ergodic averages for three polynomials and applications. *Trans. Amer. Math. Soc.*, **360** (2008), 5435–5475.
- [9] H. FURSTENBERG. Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions. *J. Anal. Math.*, **31** (1977), 204–256.
- [10] H. FURSTENBERG. *Recurrence in Ergodic Theory and Combinatorial Number Theory*. Princeton University Press, Princeton, New Jersey, 1981.
- [11] H. FURSTENBERG AND Y. KATZNELSON. An ergodic Szemerédi theorem for commuting transformations. *J. Anal. Math.*, **34** (1978), 275–291.
- [12] O. KALLENBERG, *Foundations of Modern Probability*. (2nd Ed.) Springer-Verlag, Berlin Heidelberg New York, 2002.
- [13] A. KHINTCHINE, Eine Verschärfung des Poincaréschen “Wiederkehrrsatzes”. *Compositio Math.*, **1** (1934), 177–179.
- [14] S. LANG. *Algebra*. Addison-Wesley, Reading, 1965.
- [15] V.V. NEMYTSKII AND V.V. STEPANOV. *Qualitative Theory of Differential Equations*. Princeton University Press, Princeton, New Jersey, 1960.
- [16] A.T. PATERSON. *Amenability*. Math. Surveys & Monographs 29, Amer. Math. Soc., 1988.
- [17] H. POINCARÉ. Sur le problème des trois corps et les équations de la dynamique. *Acta Math.*, **13** (1890), 1–270.
- [18] A. J. POTTS. Multiple ergodic averages for flows and an application. *Illinois J. Math.*, **55** (2011), 589–621.